#6

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to:

Assistant Commissioner for Patents
Washington, D.C. 20231

On September 21, 2001

TOWNSEND and TOWNSEND and CREW LLP

By: Joy M Marshall

PATENT
Attorney Docket No.: 02307O-099910US
Client Ref. No.: 99-404-2

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of:

FISCHER *et al.*

Application No.: 09/840,743

Filed: April 23, 2001

For: NUCLEIC ACIDS THAT CONTROL PLANT DEVELOPMENT

Examiner: Not yet assigned

Art Unit: Not yet assigned

COMMUNICATION UNDER

37 C.F.R. §§ 1.821-1.825

AND

PRELIMINARY AMENDMENT

Box SEQUENCE
Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

In response to the request to comply with Requirements for Patent Applications Containing Nucleotide Sequence and/or Amino Acid Sequence Disclosures, 37 C.F.R. §§ 1.821-1.825, that accompanied the Notice to File Missing Parts of Nonprovisional Application mailed May 23, 2001, Applicants submit herewith the required paper copy and computer readable copy of the Sequence Listing. Please amend the specification in adherence with 37 C.F.R. §§ 1.821-1.825 as follows.

In the Specification:

Please replace the paragraph beginning at page 3, line 2 with the following:

--This invention provides isolated nucleic acids comprising a polynucleotide sequence, or its complement, encoding a DMT polypeptide comprising an amino acid sequence with at least 70% sequence identity to at least one of the following consensus sequences:

DMT Domain A

KV<1>(I,l)D(D,p)(E,v)T<3>W<1>(L,v)L(M,l)(E,d)<0-2>D(K,e)<1>(K,t)<1>(K,a)(W,k)(W,l)<1>(E,k)ER<2>F<1>(G,t)R<1>(D,n)(S,l)FI(A,n)RM(H,r)<1>(V,l)QG(D,n)R<1>F<1>(P,q)WKGSVVDSV(I,v)GVFLTQN(V,t)D(H,y)(L,s)SS(S,n)A(F,y)M<1>(L,v)A(A,s)<1>FP (SEQ ID NO:71)

DMT Domain B

W(D,n)<1>(L,f)R<5>E<3-6>D(S,t)<1>(D,n)(Y,w)<3>R<10>I<2>RG(M,q)(N,f)<2>L(A,s)<1>RI<2-12>FL<3>V<2>(H,n)G<1>IDLEWLR<2>(P,d)(P,s)(D,h)<1>(A,v)K<1>(Y,f)LL(S,e)(I,f)<1>G(L,i)GLKS(V,a)ECVRLL<1>L(H,k)<2>AFPVDTNVGRI(A,c)VR(M,l)G(W,l)VPL(Q,e)PLP<2>(L,v)Q(L,m)H(L,q)L(E,f)<1>YP<1>(L,m)(E,d)(S,n)(I,v)QK(F,y)LWPRLCKL(D,p)Q<1>TLYELHY(Q,h)(L,m)ITFGK<0-2>FCTK<2>PNCNACPM(R,k)<0-2>EC(R,k)(H,y)(F,y)(A,s)SA<1> (A,v)<0-10>S(A,s)(R,k)<1>(A,l)L(P,e)<1>(P,t) (SEQ ID NO:72)

DMT Domain C.

P(I,l)(I,v)E(E,f)P<1>(S,t)P<2-5>E<0-15>(D,a)IE(D,e)<4-23>(I,v)P<1>I<1>(L,f)(N,d)<8-17>(S,a)<1>(A,d)LV<8>(I,l)P<2-5>(K,r)(L,m)K<4>LRTEH<1>V(Y,f)(E,v)LPD <1>H<1>(L,i)L(E,k)<1>(D,e)D(P,i)<2>YLL(A,s)IW(T,q)P(G,d)(E,g)<6-8>(P,s)<3>C<6-10>(M,l)C<4>C<2>C<3>(R,k)E<5>(V,f)RGT(L,i)L<0-22>(L,v)FADH<1>(S,t)(S,r)<2>PI<3>(R,t)<3>(W,k)<1>L<1>(R,k)R<4>G(T,s)(S,t)<2>(S,t)I(F,c)(R,k)(G,l)L<1>(T,v)<2>I<2>(C,n)F(W,q)<1>G(F,y)(V,l)C(V,l)R<1>F(E,d)<3>(R,g)<1>P(R,k)<1>L<2>(R,h)LH<2>(A,v)SK (SEQ ID NO:73)--

Please replace the paragraph beginning at page 11, line 3 with the following:

--A "*DMT* nucleic acid" or "*DMT* polynucleotide sequence" of the invention is a subsequence or full length polynucleotide sequence of a gene which encodes a polypeptide involved in control of reproductive development and which, when the maternal allele is mutated or when DMT activity is reduced or eliminated in a maternal tissue or plant, allows for increased production of the endosperm and/or abortion of the embryo. In addition, overexpression of DMT in plants results in delayed time to flowering. Moreover, DMT is necessary and sufficient for expression of *MEDEA* in a plant cell. An exemplary nucleic acid of the invention is the *Arabidopsis DMT* sequence (SEQ ID NO:1). Additional *DMT* nucleic acid and amino acid sequences from a variety of plant species are also provided (e.g., SEQ ID NOs: 7-70). *DMT* polynucleotides are defined by their ability to hybridize under defined conditions to the exemplified nucleic acids or PCR products derived from them. A DMT polynucleotide is typically at least about 30-40 nucleotides to about 7000, usually less than about 10,000 nucleotides in length. More preferably, DMT polynucleotides contain a coding sequence of from about 100 to about 5500 nucleotides, often from about 500 to about 3600 nucleotides in length. A DMT polypeptide is typically at least 500 amino acids, typically at least 1000 amino acids, more typically at least 1500 amino acids. In some embodiments, a DMT polypeptide comprises fewer than 2000 amino acids, more typically fewer than 3000 amino acid and still more typically fewer than 5000 or 7500 amino acid in length.--

Please replace the paragraph beginning at page 11, line 22 with the following:

--As described below, *DMT* nucleic acid sequences encode polypeptides with substantial identity to at least one of following the consensus sequences:

DMT Domain A

KV<1>(I,l)D(D,p)(E,v)T<3>W<1>(L,v)L(M,l)(E,d)<0-2>D(K,e)<1>(K,t)<1>(K,a)(W,k)(W,l)<1>(E,k)ER<2>F<1>(G,t)R<1>(D,n)(S,l)FI(A,n)RM(H,r)<1>(V,l)QG(D,n)R<1>F<1>(P,q)WKGSVVDSV(I,v)GVFLTQN(V,t)D(H,y)(L,s)SS(S,n)A(F,y)M<1>(L,v)A(A,s)<1>FP (SEQ ID NO:71)

DMT Domain B

W(D,n)<1>(L,f)R<5>E<3-6>D(S,t)<1>(D,n)(Y,w)<3>R<10>I<2>RG(M,q)(N,f)<2>L(A,s)<1>RI<2-12>FL<3>V<2>(H,n)G<1>IDLEWLR<2>(P,d)(P,s)(D,h)<1>(A,v)K<1>(Y,f)LL(S,e)(I,f)<1>G(L,i)GLKS(V,a)ECVRLL<1>L(H,k)<2>AFPVDTNVGRI(A,c)VR(M,l)G(W,l)VPL(Q,e)PLP<2>(L,v)Q(L,m)H(L,q)L(E,f)<1>YP<1>(L,m)(E,d)(S,n)(I,v)QK(F,y)LWPRLCKL(D,p)Q<1>TLYELHY(Q,h)(L,m)ITFGK<0-2>FCTK<2>PNCNACPM(R,k)<0-2>EC(R,k)(H,y)(F,y)(A,s)SA<1> (A,v)<0-10>S(A,s)(R,k)<1>(A,l)L(P,e)<1>(P,t) (SEQ ID NO:72)

DMT Domain C.

P(I,l)(I,v)E(E,f)P<1>(S,t)P<2-5>E<0-15>(D,a)IE(D,e)<4-23>(I,v)P<1>I<1>(L,f)(N,d)<8-17>(S,a)<1>(A,d)LV<8>(I,l)P<2-5>(K,r)(L,m)K<4>LRTEH<1>V(Y,f)(E,v)LPD <1>H<1>(L,i)L(E,k)<1>(D,e)D(P,i)<2>YLL(A,s)IW(T,q)P(G,d)(E,g)<6-8>(P,s)<3>C<6-10>(M,l)C<4>C<2>C<3>(R,k)E<5>(V,f)RGT(L,i)L<0-22>(L,v)FADH<1>(S,t)(S,r)<2>PI<3>(R,t)<3>(W,k)<1>L<1>(R,k)R<4>G(T,s)(S,t)<2>(S,t)I(F,c)(R,k)(G,l)L<1>(T,v)<2>I<2>(C,n)F(W,q)<1>G(F,y)(V,l)C(V,l)R<1>F(E,d)<3>(R,g)<1>P(R,k)<1>L<2>(R,h)LH<2>(A,v)SK (SEQ ID NO:73)--

Please replace the paragraph beginning at page 12, line 14 with the following:

--In addition, the following consensus sequence spanning all three domains was identified:

<9-14>(T,q)(A,i)(S,k)(I,l)<3>(A,r)(S,k)<1>(G,m)<2>(S,r)(P,k)<2>(K,f)<2>(E,l)K
<0-1>K<0-3>(P,r)<2>(P,r)<1>(K,r)(K,r)(G,d)(R,k)<1>(G,v)<1>(K,g)<3-5>(P,s)(P,k)
<3>(S,n)<1>(I,l)<0-2>(Q,d)<9>(P,q)<4>(K,a)(P,s)<14-16>(P,a)<4>L<0-10>D<1>(I,l)
<0-4>(L,n)<12-46>(K,d)<2-7>(P,a)KV<1>(I,l)D(D,p)(E,v)T<3>W<1>(L,v)L(M,l)(E,d)
<0-2>D(K,e)<1>(K,t)<1>(K,a)(W,k)(W,l)<1>(E,k)ER<2>F<1>(G,t)R<1>(D,n)(S,l)FI
(A,n)RM(H,r)<1>(V,l)QG(D,n)R<1>F<1>(P,q)WKGSVVDSV(I,v)GVFLTQN(V,t)D(H,y)(L,s)SS
(S,n)A(F,y)M<1>(L,v)A(A,s)<1>FP<0-16>(P,v)<6-15>(S,h)<3>(E,d)<10-24>(S,t)<1>
(S,e)<6>(K,n)<8-55>(E,i)<8-9>(I,v)<1>(N,s)<1-4>(E,d)<1>(E,s)<4>(Q,l)<0-11>(D,h)
<1>(F,m)<5>(Q,n)<0-3>(G,e)<2>(G,d)S<1>(K,d)<7-11>(T,m)<2>(V,l)<3>(S,q)<6-10>
(S,e)<2-3>(S,v)<19-25>(T,s)<16-28>(R,s)<2-6>(T,p)<5>(P,k)<10>(Q,e)<4>(D,s)<1-4>
(S,r)<5>(D,p)<3>(N,d)<3>(P,y)<2>(F,s)<1>(R,k)<1>(G,s)<1>(S,a)(V,r)(P,e)<3>(T,s)
<3-6>(I,l)<3>(P,e)<1>E<3-5>(L,q)<1>(G,c)<1>(S,h)(S,n)<1>(V,q)<1>(E,d)<3>T(Q,e)
<1-2>(N,g)<3>(E,n)<20-30>(N,a)(P,g)<1-6>(S,l)<25-46>(Q,d)W(D,n)<1>(L,f)R<5>E
<3-6>D(S,t)<1>(D,n)(Y,w)<3>R<10>I<2>RG(M,q)(N,f)<2>L(A,s)<1>RI<2-12>FL<3>V<2>
(H,n)G<1>IDLEWLR<2>(P,d)(P,s)(D,h)<1>(A,v)K<1>(Y,f)LL(S,e)(I,f)<1>G(L,i)GLKS
(V,a)ECVRLL<1>L(H,k)<2>AFPVDTNVGRI(A,c)VR(M,l)G(W,l)VPL(Q,e)PLP<2>(L,v)Q(L,m)H
(L,q)L(E,f)<1>YP<1>(L,m)(E,d)(S,n)(I,v)QK(F,y)LWPRLCKL(D,p)Q<1>TLYELHY(Q,h)
(L,m)ITFGK<0-2>FCTK<2>PNCNACPM(R,k)<0-2>EC(R,k)(H,y)(F,y)(A,s)SA<1>(A,v)<0-10>
S(A,s)(R,k)<1>(A,l)L(P,e)<1>(P,t)(E,q)<7-16>P(I,l)(I,v)E(E,f)P<1>(S,t)P<2-5>E
<0-15>(D,a)IE(D,e)<4-23>(I,v)P<1>I<1>(L,f)(N,d)<8-17>(S,a)<1>(A,d)LV<8>(I,l)P
<2-5>(K,r)(L,m)K<4>LRTEH<1>V(Y,f)(E,v)LPD<1>H<1>(L,i)L(E,k)<1>(D,e)D(P,i)<2>YLL
(A,s)IW(T,q)P(G,d)(E,g)<6-8>(P,s)<3>C<6-10>(M,l)C<4>C<2>C<3>(R,k)E<5>(V,f)RGT
(L,i)L<0-22>(L,v)FADH<1>(S,t)(S,r)<2>PI<3>(R,t)<3>(W,k)<1>L<1>(R,k)R<4>G(T,s)
(S,t)<2>(S,t)I(F,c)(R,k)(G,l)L<1>(T,v)<2>I<2>(C,n)F(W,q)<1>G(F,y)(V,l)C(V,l)R
<1>F(E,d)<3>(R,g)<1>P(R,k)<1>L<2>(R,h)LH<2>(A,v)SK (SEQ ID NO:74)--

Please replace the paragraph beginning at page 13, line 31 with the following:

--Amino acids 1167-1368 is related to proteins in the HhH-GPD superfamily. Amino acids 1,271 to 1,304 correspond to the conserved HhH-GPD motif. The corresponding DMT sequence is

DKAKDYLLSIRGLGLKSVECVRLLTLHNLAFPVD (SEQ ID NO:75). Secondary structure prediction (Jpred program) indicates that DMT has two alpha-helices (1,271 - 1,279 and 1,286 to 1,295) that correspond to the conserved alphaK and alphaL helices in the HhH-GPD motif of the crystallized hOGG1 DNA repair protein (Bruner et al *Nature* 403:859-866 (2000)). In between the two helices (1280 to 1285), is a hairpin with conserved glycines (G1282 and G1284). Amino acids 1286 to 1295 are related to the alphaL helix of hOGG1, which contacts the DNA backbone (Bruner *et al Nature* 403:859-866 (2000). Thus, without intending to limit the scope of the invention, it is believed this region of DMT contacts the DNA. The catalytic lysine (K1286) and aspartic acid (D1304) residues are conserved in the HhH-GPD motif of DMT. Without intending to limit the scope of the invention, by analogy to hOGG1, K1286 is predicted to displace the modified base and to promote conjugate elimination of the 3'-phosphodiester bond. Without intending to limit the scope of the invention, by analogy to hOGG1, D1304 is believed to assist the reaction by transferring protons to and from K1286.--

Please replace the paragraph beginning at page 23, line 11 with the following:

--Appropriate primers and probes for identifying *DMT* sequences from plant tissues are generated from comparisons of the sequences provided here with other related genes. For instance, *DMT* can be compared to the other endonuclease III genes, such as Genbank Accession No. AE002073. Using these techniques, one of skill can identify conserved regions in the nucleic acids disclosed here to prepare the appropriate primer and probe sequences. Primers that specifically hybridize to conserved regions in *DMT* genes can be used to amplify sequences from widely divergent plant species. Appropriate primers for amplification of the genomic region or cDNA of *DMT* include the following primers (SEQ ID NOS:76-119):

Xba-SKEN-7; CCTCTAGAGGAATTGTCGGCAAAATCGAG

SKB-8; GGAGAGACGGTTATTGTCAACC

SKB-7; AAAAGTCTACAAGGGAGAGAGAGT

SKB-5; GTAGATGTACATACGTACC

SKEN-8; GCATCCTCCAACAAGTAACAATCCACTC

SKB-6; CACTGAGATTAATTCTTCAGACTCG

SKEN-3.5; CTCAGGCGAGTCAATGCCGGAGAACAC

SKEN-3; CGAGGGCTGATCCGGGGGATAGATATTTT

SKEN-2; CCCCCGGATCAGCCCTCGAATTC

SKEN-1; CCCCTGTCTACAAATTCACCACCTGG

SKEL-4; CTGACCCAACTGCTTCTCTTC

skes1.5; TCACCTGTTCTGAACAGACTGG

SKES-1.4; CAGCAGACGAGTCCATAATGCTCTGC

SKES-2.4; GGTTTGCCTTCCACGACCACC

SKES-1; GGAAGCCACGCAAAGCTGCAACTCAGG

SKES-2.45; GAGTTGCAGCTTTGCGTGGCTTCC

SKES2.5; TTCAGACTCAGAGTCACCTTGC

SKES-2; ACCAGCAGCCTTGCTTGGCC

SKES-3; CATGCCAGAGAAGCAGGGCTCC

SKES3.5; CGATGATACTGTCTCTTCGAGC

SKES-6; CCTCCGCCTGCTCATGCCTCAG

SKEN-4; GTCCATCAGGAGAACTTCTGTGTCAGGAT

SKES-4; GGGAACAAGTGCACCATCTCC

SKEN-6; GCTCTCATAGGGAACAAGTGCACCATCTC

SKES-5; CGCTCGCATGCACCTGGTAC

SKB-1; GGAGGGAATCGAGCAGCTAGAG

SKB-2; GAGCAGCTAAGGGACTGTTCAAACTC

SKB-3; CCAGGAATGGGATTGTCCGG

3' RACE-2; CTTGGACGGCGCTTGAGGAACC

3' RACE-1; GCCTACAAGCCAGTGGGATAG

cDNA-1; GCCAAGGACTATCTCTTGAGC

SKB-4; GGATGGACTCGAGCACTGGG

SKE2.2-4; AGAGGAGAGTGCAGACACTTTG

cDNA-3; GAGGACCCTGACGAGATCCCAAC

cDNA-9; CCATGTGTTCCCGTAGAGTCATTCC

2.2+SKE-1; ATGGAGCTCCAAGAAGGTGACATG

cDNA-5; CAGAAGTGTGGAGGGAAAGCGTCTGGC

cDNA-4; CCCTCAGACTGTTACACTCAGAAC

cDNA-2; CCCGTTGAGCGGAAAACTTCCTCTCATGGC

cDNA-7; GGAAAGGATTCGTATGTGTCCGTGG

SKEN-5; GCAATGCGTTTGCTTTCTTCCAGTCATCT

cDNA-6; GAGGAGAGCAGAGAAGCAATGCGTTTGC

cDNA-8; GTTAGAGAGAAAATAAATAACCC

2.2+SKE-3; CCGTAAACAACACCGGATACAC--

Please replace the paragraph beginning at page 40, line 1 with the following:

--5'- and 3'- RACE were used to delineate the 5'- and 3'-ends of the cDNA, respectively. 5'-RACE was carried out using reagents and protocols provided by 5' RACE System for Rapid Amplification of cDNA Ends, Version 2.0, GIBCO BRL, LIFE TECHNOLOGIES, Grand Island, NY and Marathon cDNA Amplification Kit, Clontech, Palo Alto, CA. Final gene specific 5'-RACE primers were SKES-4 (GGGAACAAGTGCACCATCTCC; SEQ ID NO:97) and SKES3.5 (CGATGATACTGTCTCTTCGAGC; SEQ ID NO:95). 3'-RACE was carried out using reagents and protocols provided by Marathon cDNA Amplification Kit, Clontech, Palo Alto. Final gene-specific 3' end was obtained from cDNA library screening.--

Please replace the paragraph beginning at page 42, line 14 with the following:

--The hallmark of the superfamily of base-excision DNA repair proteins is a helix-hairpin-helix structural element followed by a Gly/Pro-rich loop and a conserved aspartic acid (i.e., HhH-GPD motif). The DMT polypeptide is 1,729 amino acids in length. Amino acids 1,271 to 1,304 correspond to the conserved HhH-GPD motif. The

DMT sequence is DKAKDYLLSIRGLGLKSVECVRLLTLHNLAFPVD (SEQ ID NO:75). The catalytic lysine (K1286) and aspartic acid (D1304) residues are conserved in the HhH-GPD motif of DMT. Secondary structure prediction (Jpred program) indicates that DMT has two alpha-helices (amino acids 1,271 - 1,279 and 1,286 to 1,295) that correspond to the conserved alphaK and alphaL helices in the HhH-GPD motif of the crystallized hOGG1 DNA repair protein (Bruner *et al Nature* 403:859-866 (2000)).--

Please replace the paragraph beginning at page 42, line 24 with the following:

--The Arabidopsis *DMT* coding sequences were also used to identify homologous sequences in both public and proprietary databases using both the BLAST and PSI-BLAST computer algorithms. This analysis revealed amino acid sequences from several plant species, including wheat, maize, rice, soybean and Arabidopsis (SEQ ID NOS:8, 9, 11, 12, 14, 15, 17, 18, 20, 22, 24, 25, 27 and 29) (SEQ ID NOs: 7-29). Based on these sequences, the following consensus sequences for DMT were determined:

DMT Domain A

KV<1>(I,l)D(D,p)(E,v)T<3>W<1>(L,v)L(M,l)(E,d)<0-2>D(K,e)<1>(K,t)<1>(K,a)(W,k)(W,l)<1>(E,k)ER<2>F<1>(G,t)R<1>(D,n)(S,l)FI(A,n)RM(H,r)<1>(V,l)QG(D,n)R<1>F<1>(P,q)WKGSVVDSV(I,v)GVFLTQN(V,t)D(H,y)(L,s)SS(S,n)A(F,y)M<1>(L,v)A(A,s)<1>FP (SEQ ID NO:71)

DMT Domain B

W(D,n)<1>(L,f)R<5>E<3-6>D(S,t)<1>(D,n)(Y,w)<3>R<10>I<2>RG(M,q)(N,f)<2>L(A,s)<1>RI<2-12>FL<3>V<2>(H,n)G<1>IDLEWLR<2>(P,d)(P,s)(D,h)<1>(A,v)K<1>(Y,f)LL(S,e)(I,f)<1>G(L,i)GLKS(V,a)ECVRLL<1>L(H,k)<2>AFPVDTNVGRI(A,c)VR(M,l)G(W,l)VPL(Q,e)PLP<2>(L,v)Q(L,m)H(L,q)L(E,f)<1>YP<1>(L,m)(E,d)(S,n)(I,v)QK(F,y)LWPRLCKL(D,p)Q<1>TLYELHY(Q,h)(L,m)ITFGK<0-2>FCTK<2>PNCNACPM(R,k)<0-2>EC(R,k)(H,y)(F,y)(A,s)SA<1> (A,v)<0-10>S(A,s)(R,k)<1>(A,l)L(P,e)<1>(P,t) (SEQ ID NO:72)

DMT Domain C.

P(I,l)(I,v)E(E,f)P<1>(S,t)P<2-5>E<0-15>(D,a)IE(D,e)<4-23>(I,v)P<1>
I<1>(L,f)(N,d)<8-17>(S,a)<1>(A,d)LV<8>(I,l)P<2-5>(K,r)(L,m)K<4>LRTEH<1>V(Y,f)
(E,v)LPD <1>H<1>(L,i)L(E,k)<1>(D,e)D(P,i)<2>YLL(A,s)IW(T,q)P(G,d)(E,g)<6-8>
(P,s)<3>C<6-10>(M,l)C<4>C<2>C<3>(R,k)E<5>(V,f)RGT(L,i)L<0-22>(L,v)FADH<1>(S,t)
(S,r)<2>PI<3>(R,t)<3>(W,k)<1>L<1>(R,k)R<4>G(T,s)(S,t)<2>(S,t)I(F,c)(R,k)(G,l)L
<1>(T,v)<2>I<2>(C,n)F(W,q)<1>G(F,y)(V,l)C(V,l)R<1>F(E,d)<3>(R,g)<1>P(R,k)<1>L
<2>(R,h)LH<2>(A,v)SK (SEQ ID NO:73)

The first consensus sequence listed above corresponds to amino acid positions 586 through 937 of SEQ ID NO:2. The second consensus sequence listed above corresponds to amino acid positions 1117 through 1722 of SEQ ID NO:2. The consensus sequence provides amino acid sequences by position using single letter amino acid abbreviations. Numbers in carrots ("<" or ">") refer to amino acid positions where there is no consensus and which therefore, can be any amino acid. Amino acid abbreviations in parentheses indicate alternative amino acids at the same position. Capitalized letters refer to predominant consensus amino acids and lower case letters refer to amino acids that are commonly found in DMT sequences, but are not predominant.--

Please cancel the present informal Sequence Listing, pages 46-100, and insert therefor the accompanying paper copy of the formal Sequence Listing, page numbers 1 to 139, at the end of the application. Cancel the page numbers for the Claims and Abstract and renumber as pages 46-50, accordingly.

REMARKS

The amendments to paragraphs beginning on pages 3, 11, 12and 42 insert SEQ ID NOS: at their appropriate locations. In addition, the sequences of DMT Domains A, B and C and the consensus sequence spanning all three domains have inserted line breaks at appropriate locations, but have not changed the respective order of amino acid residues.

The amendment to the paragraph beginning on page 42, line 24, at lines 27 and 28, changes those assigned sequence identifiers to reflect only those SEQ ID NOS: which refer to "amino acid sequences", from line 26.

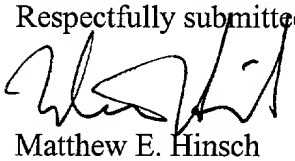
Applicants request entry of this amendment in adherence with 37 C.F.R. §§1.821 to 1.825. This amendment is accompanied by a floppy disk containing the above named sequences, SEQ ID NOS:1-119, in computer readable form, and a paper copy of the sequence information which has been printed from the floppy disk.

The information contained in the computer readable disk was prepared through the use of the software program "PatentIn" and is identical to that of the paper copy. This amendment contains no new matter.

Attached hereto is a marked-up version of the changes made to the Specification by the current Amendment. The attached pages are captioned **"VERSION WITH MARKINGS TO SHOW CHANGES MADE."**

If the Examiner believes a telephone conference would expedite prosecution of this application, please telephone the undersigned at 415-576-0200.

Respectfully submitted,



Matthew E. Hinsch
Reg. No. 47,651

TOWNSEND and TOWNSEND and CREW LLP
Two Embarcadero Center, 8th Floor
San Francisco, California 94111-3834
Tel: (415) 576-0200
Fax: (415) 576-0300
MEH:dmw

VERSION WITH MARKINGS TO SHOW CHANGES MADE

In the Specification:

Paragraph beginning at line 2 of page 3 has been amended as follows:

This invention provides isolated nucleic acids comprising a polynucleotide sequence, or its complement, encoding a DMT polypeptide comprising an amino acid sequence with at least 70% sequence identity to at least one of the following consensus sequences:

DMT Domain A

KV<1>(I,l)D(D,p)(E,v)T<3>W<1>(L,v)L(M,l)(E,d)<0-2>D(K,e)<1>(K,t)<1>(K,a)(W,k)(W,l)<1>(E,k)ER<2>F<1>(G,t)R<1>(D,n)(S,l)FI(A,n)RM(H,r)<1>(V,l)QG(D,n)R<1>F<1>(P,q)WKGSVVDSV(I,v)GVFLTQN(V,t)D(H,y)(L,s)SS(S,n)A(F,y)M<1>(L,v)A(A,s)<1>FP (SEQ ID NO:71)

DMT Domain B

W(D,n)<1>(L,f)R<5>E<3-6>D(S,t)<1>(D,n)(Y,w)<3>R<10>I<2>RG(M,q)(N,f)<2>L(A,s)<1>RI<2-12>FL<3>V<2>(H,n)G<1>IDLEWLR<2>(P,d)(P,s)(D,h)<1>(A,v)K<1>(Y,f)LL(S,e)(I,f)<1>G(L,i)GLKS(V,a)ECVRLL<1>L(H,k)<2>AFPVDTNVGRI(A,c)VR(M,l)G(W,l)VPL(Q,e)PLP<2>(L,v)Q(L,m)H(L,q)L(E,f)<1>YP<1>(L,m)(E,d)(S,n)(I,v)QK(F,y)LWPRLCKL(D,p)Q<1>TLYELHY(Q,h)(L,m)ITFGK<0-2>FCTK<2>PNCNACPM(R,k)<0-2>EC(R,k)(H,y)(F,y)(A,s)SA<1> (A,v)<0-10>S(A,s)(R,k)<1>(A,l)L(P,e)<1>(P,t) (SEQ ID NO:72)

DMT Domain C.

P(I,l)(I,v)E(E,f)P<1>(S,t)P<2-5>E<0-15>(D,a)IE(D,e)<4-23>(I,v)P<1>I<1>(L,f)(N,d)<8-17>(S,a)<1>(A,d)LV<8>(I,l)P<2-5>(K,r)(L,m)K<4>LRTEH<1>V(Y,f)(E,v)LPD <1>H<1>(L,i)L(E,k)<1>(D,e)D(P,i)<2>YLL(A,s)IW(T,q)P(G,d)(E,g)<6-8>(P,s)<3>C<6-10>(M,l)C<4>C<2>C<3>(R,k)E<5>(V,f)RGT(L,i)L<0-22>(L,v)FADH<1>(S,t)(S,r)<2>PI<3>(R,t)<3>(W,k)<1>L<1>(R,k)R<4>G(T,s)(S,t)<2>(S,t)I(F,c)(R,k)(G,l)L<1>(T,v)<2>I<2>(C,n)F(W,q)<1>G(F,y)(V,l)C(V,l)R<1>F(E,d)<3>(R,g)<1>P(R,k)<1>L<2>(R,h)LH<2>(A,v)SK (SEQ ID NO:73)

Paragraph beginning at line 3 of page 11 has been amended as follows:

A "*DMT* nucleic acid" or "*DMT* polynucleotide sequence" of the invention is a subsequence or full length polynucleotide sequence of a gene which encodes a polypeptide involved in control of reproductive development and which, when the maternal allele is mutated or when DMT activity is reduced or eliminated in a maternal tissue or plant, allows for increased production of the endosperm and/or abortion of the embryo. In addition, overexpression of DMT in plants results in delayed time to flowering. Moreover, DMT is necessary and sufficient for expression of *MEDEA* in a plant cell. An exemplary nucleic acid of the invention is the *Arabidopsis DMT* sequence (SEQ ID NO:1). Additional *DMT* nucleic acid and amino acid sequences from a variety of plant species are also provided (e.g., SEQ ID NOs: 7-70). *DMT* polynucleotides are defined by their ability to hybridize under defined conditions to the exemplified nucleic acids or PCR products derived from them. A DMT polynucleotide is typically at least about 30-40 nucleotides to about 7000, usually less than about 10,000 nucleotides in length. More preferably, DMT polynucleotides contain a coding sequence of from about 100 to about 5500 nucleotides, often from about 500 to about 3600 nucleotides in length. A DMT polypeptide is typically at least 500 amino acids, typically at least 1000 amino acids, more typically at least 1500 amino acids. In some embodiments, a DMT polypeptide comprises fewer than 2000 amino acids, more typically fewer than 3000 amino acid and still more typically fewer than 5000 or 7500 amino acid in length.

Paragraph beginning at line 22 of page 11 has been amended as follows:

As described below, *DMT* nucleic acid sequences encode polypeptides with substantial identity to at least one of following the consensus sequences:

DMT Domain A

KV<1>(I,l)D(D,p)(E,v)T<3>W<1>(L,v)L(M,l)(E,d)<0-2>D(K,e)<1>(K,t)<1>(K,a)(W,k)(W,l)<1>(E,k)ER<2>F<1>(G,t)R<1>(D,n)(S,l)FI(A,n)RM(H,r)<1>(V,l)QG(D,n)R<1>F<1>(P,q)WKGSVVDSV(I,v)GVFLTQN(V,t)D(H,y)(L,s)SS(S,n)A(F,y)M<1>(L,v)A(A,s)<1>FP (SEQ ID NO:71)

DMT Domain B

W(D,n)<1>(L,f)R<5>E<3-6>D(S,t)<1>(D,n)(Y,w)<3>R<10>I<2>RG(M,q)(N,f)<2>L(A,s)<1>RI<2-12>FL<3>V<2>(H,n)G<1>IDLEWLR<2>(P,d)(P,s)(D,h)<1>(A,v)K<1>(Y,f)LL(S,e)(I,f)<1>G(L,i)GLKS(V,a)ECVRLL<1>L(H,k)<2>AFPVDTNVGRI(A,c)VR(M,l)G(W,l)VPL(Q,e)PLP<2>(L,v)Q(L,m)H(L,q)L(E,f)<1>YP<1>(L,m)(E,d)(S,n)(I,v)QK(F,y)LWPRLCKL(D,p)Q<1>TLYELHY(Q,h)(L,m)ITFGK<0-2>FCTK<2>PNCNACPM(R,k)<0-2>EC(R,k)(H,y)(F,y)(A,s)SA<1> (A,v)<0-10>S(A,s)(R,k)<1>(A,l)L(P,e)<1>(P,t) (SEQ ID NO:72)

DMT Domain C.

P(I,l)(I,v)E(E,f)P<1>(S,t)P<2-5>E<0-15>(D,a)IE(D,e)<4-23>(I,v)P<1>I<1>(L,f)(N,d)<8-17>(S,a)<1>(A,d)LV<8>(I,l)P<2-5>(K,r)(L,m)K<4>LRTEH<1>V(Y,f)(E,v)LPD <1>H<1>(L,i)L(E,k)<1>(D,e)D(P,i)<2>YLL(A,s)IW(T,q)P(G,d)(E,g)<6-8>(P,s)<3>C<6-10>(M,l)C<4>C<2>C<3>(R,k)E<5>(V,f)RGT(L,i)L<0-22>(L,v)FADH<1>(S,t)(S,r)<2>PI<3>(R,t)<3>(W,k)<1>L<1>(R,k)R<4>G(T,s)(S,t)<2>(S,t)I(F,c)(R,k)(G,l)L<1>(T,v)<2>I<2>(C,n)F(W,q)<1>G(F,y)(V,l)C(V,l)R<1>F(E,d)<3>(R,g)<1>P(R,k)<1>L<2>(R,h)LH<2>(A,v)SK (SEQ ID NO:73)

Paragraph beginning at line 14 of page 12 has been amended as follows:

In addition, the following consensus ~~eansensus~~ sequence spanning all three domains was ~~were~~ identified:

<9-14>(T,q)(A,i)(S,k)(I,l)<3>(A,r)(S,k)<1>(G,m)<2>(S,r)(P,k)<2>(K,f)<2>(E,l)K
<0-1>K<0-3>(P,r)<2>(P,r)<1>(K,r)(K,r)(G,d)(R,k)<1>(G,v)<1>(K,g)<3-5>(P,s)(P,k)
<3>(S,n)<1>(I,l)<0-2>(Q,d)<9>(P,q)<4>(K,a)(P,s)<14-16>(P,a)<4>L<0-10>D<1>(I,l)
<0-4>(L,n)<12-46>(K,d)<2-7>(P,a)KV<1>(I,l)D(D,p)(E,v)T<3>W<1>(L,v)L(M,l)(E,d)
<0-2>D(K,e)<1>(K,t)<1>(K,a)(W,k)(W,l)<1>(E,k)ER<2>F<1>(G,t)R<1>(D,n)(S,l)FI
(A,n)RM(H,r)<1>(V,l)QG(D,n)R<1>F<1>(P,q)WKGSVVDSV(I,v)GVFLTQN(V,t)D(H,y)(L,s)SS
(S,n)A(F,y)M<1>(L,v)A(A,s)<1>FP<0-16>(P,v)<6-15>(S,h)<3>(E,d)<10-24>(S,t)<1>
(S,e)<6>(K,n)<8-55>(E,i)<8-9>(I,v)<1>(N,s)<1-4>(E,d)<1>(E,s)<4>(Q,l)<0-11>(D,h)
<1>(F,m)<5>(Q,n)<0-3>(G,e)<2>(G,d)S<1>(K,d)<7-11>(T,m)<2>(V,l)<3>(S,q)<6-10>
(S,e)<2-3>(S,v)<19-25>(T,s)<16-28>(R,s)<2-6>(T,p)<5>(P,k)<10>(Q,e)<4>(D,s)<1-4>
(S,r)<5>(D,p)<3>(N,d)<3>(P,y)<2>(F,s)<1>(R,k)<1>(G,s)<1>(S,a)(V,r)(P,e)<3>(T,s)
<3-6>(I,l)<3>(P,e)<1>E<3-5>(L,q)<1>(G,c)<1>(S,h)(S,n)<1>(V,q)<1>(E,d)<3>T(Q,e)
<1-2>(N,g)<3>(E,n)<20-30>(N,a)(P,g)<1-6>(S,l)<25-46>(Q,d)W(D,n)<1>(L,f)R<5>E
<3-6>D(S,t)<1>(D,n)(Y,w)<3>R<10>I<2>RG(M,q)(N,f)<2>L(A,s)<1>RI<2-12>FL<3>V<2>
(H,n)G<1>IDLEWLR<2>(P,d)(P,s)(D,h)<1>(A,v)K<1>(Y,f)LL(S,e)(I,f)<1>G(L,i)GLKS
(V,a)ECVRLL<1>L(H,k)<2>AFPVDTNVGRI(A,c)VR(M,l)G(W,l)VPL(Q,e)PLP<2>(L,v)Q(L,m)H
(L,q)L(E,f)<1>YP<1>(L,m)(E,d)(S,n)(I,v)QK(F,y)LWPRLCKL(D,p)Q<1>TLYELHY(Q,h)
(L,m)ITFGK<0-2>FCTK<2>PNCNACPM(R,k)<0-2>EC(R,k)(H,y)(F,y)(A,s)SA<1>(A,v)<0-10>
S(A,s)(R,k)<1>(A,l)L(P,e)<1>(P,t)(E,q)<7-16>P(I,l)(I,v)E(E,f)P<1>(S,t)P<2-5>E
<0-15>(D,a)IE(D,e)<4-23>(I,v)P<1>I<1>(L,f)(N,d)<8-17>(S,a)<1>(A,d)LV<8>(I,l)P
<2-5>(K,r)(L,m)K<4>LRTEH<1>V(Y,f)(E,v)LPD<1>H<1>(L,i)L(E,k)<1>(D,e)D(P,i)<2>YLL
(A,s)IW(T,q)P(G,d)(E,g)<6-8>(P,s)<3>C<6-10>(M,l)C<4>C<2>C<3>(R,k)E<5>(V,f)RGT
(L,i)L<0-22>(L,v)FADH<1>(S,t)(S,r)<2>PI<3>(R,t)<3>(W,k)<1>L<1>(R,k)R<4>G(T,s)
(S,t)<2>(S,t)I(F,c)(R,k)(G,l)L<1>(T,v)<2>I<2>(C,n)F(W,q)<1>G(F,y)(V,l)C(V,l)R
<1>F(E,d)<3>(R,g)<1>P(R,k)<1>L<2>(R,h)LH<2>(A,v)SK (SEQ ID NO:74)

Paragraph beginning at line 31 of page 13 has been amended as follows:

Amino acids 1167-1368 is related to proteins in the HhH-GPD superfamily. Amino acids 1,271 to 1,304 correspond to the conserved HhH-GPD motif. The corresponding DMT sequence is

DKAKDYLLSIRGLGLKSVECVRLLTLHNLAFPVD (SEQ ID NO:75). Secondary structure prediction (Jpred program) indicates that DMT has two alpha-helices (1,271 - 1,279 and 1,286 to 1,295) that correspond to the conserved alphaK and alphaL helices in the HhH-GPD motif of the crystallized hOGG1 DNA repair protein (Bruner et al *Nature* 403:859-866 (2000)). In between the two helices (1280 to 1285), is a hairpin with conserved glycines (G1282 and G1284). Amino acids 1286 to 1295 are related to the alphaL helix of hOGG1, which contacts the DNA backbone (Bruner *et al Nature* 403:859-866 (2000). Thus, without intending to limit the scope of the invention, it is believed this region of DMT contacts the DNA. The catalytic lysine (K1286) and aspartic acid (D1304) residues are conserved in the HhH-GPD motif of DMT. Without intending to limit the scope of the invention, by analogy to hOGG1, K1286 is predicted to displace the modified base and to promote conjugate elimination of the 3'-phosphodiester bond. Without intending to limit the scope of the invention, by analogy to hOGG1, D1304 is believed to assist the reaction by transferring protons to and from K1286.

Paragraph beginning at line 11 of page 23 has been amended as follows:

Appropriate primers and probes for identifying *DMT* sequences from plant tissues are generated from comparisons of the sequences provided here with other related genes. For instance, *DMT* can be compared to the other endonuclease III genes, such as Genbank Accession No. AE002073. Using these techniques, one of skill can identify conserved regions in the nucleic acids disclosed here to prepare the appropriate primer and probe sequences. Primers that specifically hybridize to conserved regions in *DMT* genes can be used to amplify sequences from widely divergent plant species. Appropriate primers for amplification of the genomic region or cDNA of *DMT* include the following primers (SEQ ID NOS:76-119):

Xba-SKEN-7; CCTCTAGAGGAATTGTCGGCAAAATCGAG

SKB-8; GGAGAGACGGTTATTGTCAACC

SKB-7; AAAAGTCTACAAGGGAGAGAGAGT

SKB-5; GTAGATGTACATACGTACC

SKEN-8; GCATCCTCCAACAAGTAACAATCCACTC

SKB-6; CACTGAGATTAATTCTTCAGACTCG

SKEN-3.5; CTCAGGCGAGTCAATGCCGGAGAACAC

SKEN-3; CGAGGGCTGATCCGGGGGATAGATATTTT

SKEN-2; CCCCCGGATCAGCCCTCGAATTC

SKEN-1; CCCCTGTCTACAAATTCACCACCTGG

SKEL-4; CTGACCCAACTGCTTCTCTTC

skes1.5; TCACCTGTTCTGAACAGACTGG

SKES-1.4; CAGCAGACGAGTCCATAATGCTCTGC

SKES-2.4; GGTTTGCCTTCCACGACCACC

SKES-1; GGAAGCCACGCAAAGCTGCAACTCAGG

SKES-2.45; GAGTTGCAGCTTTGCGTGGCTTCC

SKES2.5; TTCAGACTCAGAGTCACCTTGC

SKES-2; ACCAGCAGCCTTGCTTGGCC

SKES-3; CATGCCAGAGAAGCAGGGCTCC

SKES3.5; CGATGATACTGTCTCTTCGAGC

SKES-6; CCTCCGCCTGCTCATGCCTCAG

SKEN-4; GTCCATCAGGAGAACTTCTGTGTCAGGAT

SKES-4; GGGAACAAGTGCACCATCTCC

SKEN-6; GCTCTCATAGGGAACAAGTGCACCATCTC

SKES-5; CGCTCGCATGCACCTGGTAC

SKB-1; GGAGGGAATCGAGCAGCTAGAG

SKB-2; GAGCAGCTAAGGGACTGTTCAAACTC

SKB-3; CCAGGAATGGGATTGTCCGG

3' RACE-2; CTTGGACGGCGCTTGAGGAACC

3' RACE-1; GCCTACAAGCCAGTGGGATAG

cDNA-1; GCCAAGGACTATCTCTTGAGC

SKB-4; GGATGGACTCGAGCACTGGG

SKE2.2-4; AGAGGAGAGTGCAGACACTTTG

cDNA-3; GAGGACCCTGACGAGATCCCAAC

cDNA-9; CCATGTGTTCCCGTAGAGTCATTCC

2.2+SKE-1; ATGGAGCTCCAAGAAGGTGACATG

cDNA-5; CAGAAGTGTGGAGGGAAAGCGTCTGGC

cDNA-4; CCCTCAGACTGTTACACTCAGAAC

cDNA-2; CCCGTTGAGCGGAAAACTTCCTCTCATGGC

cDNA-7; GGAAAGGATTCGTATGTGTCCGTGG

SKEN-5; GCAATGCGTTTGCTTTCTTCCAGTCATCT

cDNA-6; GAGGAGAGCAGAGAAGCAATGCGTTTGC

cDNA-8; GTTAGAGAGAAAATAAATAACCC

2.2+SKE-3; CCGTAAACAACACCGGATACAC

Paragraph beginning at line 1 of page 40 has been amended as follows:

5'- and 3'- RACE were used to delineate the 5'- and 3'-ends of the cDNA, respectively. 5'-RACE was carried out using reagents and protocols provided by 5' RACE System for Rapid Amplification of cDNA Ends, Version 2.0, GIBCO BRL, LIFE TECHNOLOGIES, Grand Island, NY and Marathon cDNA Amplification Kit, Clontech, Palo Alto, CA. Final gene specific 5'-RACE primers were SKES-4 (GGGAACAAGTGCACCATCTCC; SEQ ID NO:97) and SKES3.5 (CGATGATACTGTCTCTTCGAGC; SEQ ID NO:95). 3'-RACE was carried out using reagents and protocols provided by Marathon cDNA Amplification Kit, Clontech, Palo Alto. Final gene-specific 3' end was obtained from cDNA library screening.

Paragraph beginning at line 14 of page 42 has been amended as follows:

The hallmark of the superfamily of base-excision DNA repair proteins is a helix-hairpin-helix structural element followed by a Gly/Pro-rich loop and a conserved aspartic acid (i.e., HhH-GPD motif). The DMT polypeptide is 1,729 amino acids in length. Amino acids 1,271 to 1,304 correspond to the conserved HhH-GPD motif. The

DMT sequence is DKAKDYLLSIRGLGLKSVECVRLLTLHNLAFPVD (SEQ ID NO:75). The catalytic lysine (K1286) and aspartic acid (D1304) residues are conserved in the HhH-GPD motif of DMT. Secondary structure prediction (Jpred program) indicates that DMT has two alpha-helices (amino acids 1,271 - 1,279 and 1,286 to 1,295) that correspond to the conserved alphaK and alphaL helices in the HhH-GPD motif of the crystallized hOGG1 DNA repair protein (Bruner *et al Nature* 403:859-866 (2000)).

Paragraph beginning at line 24 of page 42 has been amended as follows:

The Arabidopsis *DMT* coding sequences were also used to identify homologous sequences in both public and proprietary databases using both the BLAST and PSI-BLAST computer algorithms. This analysis revealed amino acid sequences from several plant species, including wheat, maize, rice, soybean and Arabidopsis (SEQ ID NOS:8, 9, 11, 12, 14, 15, 17, 18, 20, 22, 24, 25, 27 and 29) ~~(SEQ ID NOs: 7-29)~~. Based on these sequences, the following consensus sequences for DMT were determined:

DMT Domain A

KV<1>(I,l)D(D,p)(E,v)T<3>W<1>(L,v)L(M,l)(E,d)<0-2>D(K,e)<1>(K,t)<1>(K,a)(W,k)(W,l)<1>(E,k)ER<2>F<1>(G,t)R<1>(D,n)(S,l)FI(A,n)RM(H,r)<1>(V,l)QG(D,n)R<1>F<1>(P,q)WKGSVVDSV(I,v)GVFLTQN(V,t)D(H,y)(L,s)SS(S,n)A(F,y)M<1>(L,v)A(A,s)<1>FP (SEQ ID NO:71)

DMT Domain B

W(D,n)<1>(L,f)R<5>E<3-6>D(S,t)<1>(D,n)(Y,w)<3>R<10>I<2>RG(M,q)(N,f)<2>L(A,s)<1>RI<2-12>FL<3>V<2>(H,n)G<1>IDLEWLR<2>(P,d)(P,s)(D,h)<1>(A,v)K<1>(Y,f)LL(S,e)(I,f)<1>G(L,i)GLKS(V,a)ECVRLL<1>L(H,k)<2>AFPVDTNVGRI(A,c)VR(M,l)G(W,l)VPL(Q,e)PLP<2>(L,v)Q(L,m)H(L,q)L(E,f)<1>YP<1>(L,m)(E,d)(S,n)(I,v)QK(F,y)LWPRLCKL(D,p)Q<1>TLYELHY(Q,h)(L,m)ITFGK<0-2>FCTK<2>PNCNACPM(R,k)<0-2>EC(R,k)(H,y)(F,y)(A,s)SA<1> (A,v)<0-10>S(A,s)(R,k)<1>(A,l)L(P,e)<1>(P,t) (SEQ ID NO:72)

DMT Domain C.

P(I,l)(I,v)E(E,f)P<1>(S,t)P<2-5>E<0-15>(D,a)IE(D,e)<4-23>(I,v)P<1>I<1>(L,f)(N,d)<8-17>(S,a)<1>(A,d)LV<8>(I,l)P<2-5>(K,r)(L,m)K<4>LRTEH<1>V(Y,f)(E,v)LPD <1>H<1>(L,i)L(E,k)<1>(D,e)D(P,i)<2>YLL(A,s)IW(T,q)P(G,d)(E,g)<6-8>(P,s)<3>C<6-10>(M,l)C<4>C<2>C<3>(R,k)E<5>(V,f)RGT(L,i)L<0-22>(L,v)FADH<1>(S,t)(S,r)<2>PI<3>(R,t)<3>(W,k)<1>L<1>(R,k)R<4>G(T,s)(S,t)<2>(S,t)I(F,c)(R,k)(G,l)L<1>(T,v)<2>I<2>(C,n)F(W,q)<1>G(F,y)(V,l)C(V,l)R<1>F(E,d)<3>(R,g)<1>P(R,k)<1>L<2>(R,h)LH<2>(A,v)SK (SEQ ID NO:73)

The first consensus sequence listed above corresponds to amino acid positions 586 through 937 of SEQ ID NO:2. The second consensus sequence listed above corresponds to amino acid positions 1117 through 1722 of SEQ ID NO:2. The consensus sequence provides amino acid sequences by position using single letter amino acid abbreviations. Numbers in carrots ("<" or ">") refer to amino acid positions where there is no consensus and which therefore, can be any amino acid. Amino acid abbreviations in parentheses indicate alternative amino acids at the same position. Capitalized letters refer to predominant consensus amino acids and lower case letters refer to amino acids that are commonly found in DMT sequences, but are not predominant.